V. COMPUTATIONAL BME FOR SOFT DATA OF PROBABILISTIC TYPE

The Bayesian Maximum Entropy (BME; Christakos, 1990, 1992) method offers considerable flexibility in the type of soft data it can process. In this Chapter I consider soft data of probabilistic type. While probabilistic type is more general than the interval type of the previous chapter, it leads to a formulation that has a higher numerical complexity, hence it is treated separately. In this chapter I propose a formulation that will lead to an efficient numerical implementation, and I will then show how this formulation is implemented numerically by taking advantage of well-suited numerical libraries. Several synthetic case studies are presented showing the superior accuracy of the BME approach compared to existing kriging methods, and valuable insights will be gained from the Equus Beds case study in Kansas.

5.1. The Processing Operator for Specificatory Knowledge of Probabilistic Type

As usual let $\boldsymbol{\chi}_{\text{hard}}$, $\boldsymbol{\chi}_{\text{soft}}$ and χ_k be the value of the Space/Time Random Field (S/TRF) at the hard data points, soft data points, and the estimation point, respectively, and we let $\boldsymbol{\chi}_{\text{map}}^{\text{T}} = [\boldsymbol{\chi}_{\text{hard}}^{\text{T}} \boldsymbol{\chi}_{\text{soft}}^{\text{T}} \chi_k]$ and $\boldsymbol{\chi}_{\text{data}}^{\text{T}} = [\boldsymbol{\chi}_{\text{hard}}^{\text{T}} \boldsymbol{\chi}_{\text{soft}}^{\text{T}}]$. In Chapter 3. we mentioned that when the specificatory knowledge consists of the hard data $\boldsymbol{\chi}_{\text{hard}}$ given by Eq. (2.25) and of the soft data $\boldsymbol{\chi}_{\text{soft}}$ of probabilistic type given by Eq. (2.27), the Y_S-operator (which processes specificatory knowledge) is given by Eq. (3.4). Writing Eq. (3.4) in terms of the prior pdf $f_G(\boldsymbol{\chi}_{\text{map}}) = Z^{-1}\exp[Y_G(\boldsymbol{\chi}_{\text{map}})]$ and combining with (3.1), we obtain the following form for the posterior BME pdf (Christakos, 1990, 1992, 1998)

$$f_K(\chi_k) = A^{-1} \int d\boldsymbol{\chi}_{\text{soft}}\ f_S(\boldsymbol{\chi}_{\text{soft}})\, f_G(\boldsymbol{\chi}_{\text{map}}) \tag{5.1}$$

where the normalization constant is expressed as $A = \int d\boldsymbol{\chi}_{\text{soft}}\ f_S(\boldsymbol{\chi}_{\text{soft}})\, f_G(\boldsymbol{\chi}_{\text{data}})$, with $f_G(\boldsymbol{\chi}_{\text{data}}) = \int d\chi_{\text{k}}\, f_G(\boldsymbol{\chi}_{\text{map}})$. Let us consider the following example to illustrate Eq. (5.1).

EXAMPLE 5.1: Assume that there is one estimation point, one hard data point and one soft data point. Let x_k be the random variable to estimate, let the random variable at the hard data point be x_h with corresponding hard (exact) measurement given by $P(x_h = \chi_h) = 1$, and let the random variable at the soft data point be x_s with corresponding soft data knowledge given by $f_S(\chi_s) = f_1 \delta_{l_1 \le \chi_s < u_1} + f_2 \delta_{u_1 = l_2 \le \chi_s < u_2}$, which means that the probability $P[l_1 \le x_s < u_1]$ is $p_1 = P[l_1 \le x_s < u_1] = (u_1 - l_1) f_1$, and the probability $P[u_1 = l_2 \le x_s < u_2]$ is $p_2 = P[l_2 \le x_s < u_2] = (u_2 - l_2) f_2$. This represents a case of soft data of the interval type, therefore the posterior BME pdf describing the random variable x_k is given in general by Eq. (5.1), which we can write here as $f_K(\chi_k) = A^{-1} \int d\chi_s\, f_S(\chi_s)\, f_G(\chi_k, \chi_h, \chi_s)$, substituting $f_S(\chi_s)$with its expression; $f_K(\chi_k) = A^{-1} (f_1 \int_{l_1}^{u_1} d\chi_s\, f_G(\boldsymbol{\chi}_{\text{map}}) + f_2 \int_{l_2 = u_1}^{u_2} d\chi_s\, f_G(\boldsymbol{\chi}_{\text{map}}))$.

A proof for Eq. (5.1) an a detailed explanation of the underlying epistemological principles is given by Christakos (1998, 1999). The processing rule of Eq. (5.1) allows to incorporate a large class of uncertain information, which was found to be very useful in practice (Serre and Christakos; 1999), and produces spatiotemporal maps that are much more accurate than those produced with existing kriging methods (Serre *et al.*; 1998). The BME method may be computationally intensive in general due to its flexibility, hence in order to produce a useful code I propose in the following a formulation which accounts for general knowledge consisting of the mean and covariance function, and lead to an efficient numerical implementation.

5.2. A proposed Formulation of the Posterior PDF for Efficient Computation

Let's assume that the Space/Time Random Field $X(\boldsymbol{p})$ has a known mean $m_x(\boldsymbol{p}) = \overline{X(\boldsymbol{p})}$, as well as a known covariance function $c_x(\boldsymbol{p}, \boldsymbol{p}')$, usually obtained from fitting to experimental data. As usual χ_k is the value of the S/TRF at the estimation point $\boldsymbol{p}_k$, $\boldsymbol{\chi}_{\text{hard}} = [\chi_1 \cdots \chi_{m_h}]^T$ are the hard data at points $\boldsymbol{p}_i$ $(i = 1, \ldots, m_h)$, and $\boldsymbol{\chi}_{\text{soft}} = [\chi_{m_h+1} \cdots \chi_m]^T$ at points $\boldsymbol{p}_i$ $(i = m_h + 1, \ldots, m)$ are soft data of the probabilistic type defined in Eq. (2.27),i.e. $\boldsymbol{\chi}_{\text{soft}}$: $P_S(\boldsymbol{x}_{\text{soft}} \le \boldsymbol{\xi}) = \int_{-\infty}^{\xi} d\boldsymbol{\chi}_{\text{soft}}\, f_S(\boldsymbol{\chi}_{\text{soft}})$. The covariance matrix $\boldsymbol{C}_{\text{map}} = \overline{(\boldsymbol{x}_{\text{map}} - \boldsymbol{m}_{\text{map}})(\boldsymbol{x}_{\text{map}} - \boldsymbol{m}_{\text{map}})^{\text{T}}}$ associated with the vector of random variable $\boldsymbol{x}_{\text{map}}$ is given by Eq. (4.3), and $\phi(\boldsymbol{x}; \bar{\boldsymbol{x}}, \boldsymbol{C})$ denote the n-point Gaussian pdf of the random vector $\boldsymbol{x}$ with mean $\bar{\boldsymbol{x}}$ and covariance matrix $\boldsymbol{C}$ as defined in Eq. (4.4). Assuming, without loss of generality, that $\boldsymbol{m}_{\text{map}} = \boldsymbol{0}$, we find that the prior pdf of Eq. (3.22) may be written as $f_G(\boldsymbol{\chi}_{\text{map}}) = \phi(\boldsymbol{\chi}_{\text{map}}; \boldsymbol{0}, \boldsymbol{C}_{\text{map}})$. In the following, it is convenient to define the partitioned matrices $\boldsymbol{C}_{h,h}$, $\boldsymbol{C}_{s,h}$, ... , $\boldsymbol{C}_{kh,kh}$ as expressed in Eqs. (4.5) and (4.6).

The posterior pdf is obtained by inserting the prior pdf $f_G(\boldsymbol{\chi}_{\text{map}}) = \phi(\boldsymbol{\chi}_{\text{map}}; \boldsymbol{0}, \boldsymbol{C}_{\text{map}})$ into Eq. (5.1). Using properties of the multivariate Gaussian pdf (see Appendices C and D), we obtain the following convenient formulation

$$f_K(\chi_k) = A^{-1}\, \phi(\chi_k; \boldsymbol{B}_{k|h} \boldsymbol{\chi}_{\text{hard}}, C_{k|h}) \int d\boldsymbol{\chi}_{\text{soft}}\, f_S(\boldsymbol{\chi}_{\text{soft}})\, \phi(\boldsymbol{\chi}_{\text{soft}}; \boldsymbol{B}_{s|kh} \boldsymbol{\chi}_{kh}, \boldsymbol{C}_{s|kh}), \tag{5.2}$$

where $\boldsymbol{\chi}_{kh}^{\text{T}} = [\chi_k\, \boldsymbol{\chi}_{\text{hard}}^T]$, $\boldsymbol{B}_{k|h} = \boldsymbol{C}_{k,h}\, \boldsymbol{C}_{h,h}^{-1}$, $C_{k|h} = C_{k,k} - \boldsymbol{B}_{k|h}\, \boldsymbol{C}_{h,k}$, $\boldsymbol{B}_{s|kh} = \boldsymbol{C}_{s,kh}\, \boldsymbol{C}_{kh,kh}^{-1}$, $\boldsymbol{C}_{s|kh} = \boldsymbol{C}_{s,s} - \boldsymbol{B}_{s|kh}\, \boldsymbol{C}_{kh,s}$, and $A' = \int d\boldsymbol{\chi}_{\text{soft}}\, f_S(\boldsymbol{\chi}_{\text{soft}})\, \phi(\boldsymbol{\chi}_{\text{soft}}; \boldsymbol{B}_{s|h} \boldsymbol{\chi}_{\text{hard}}, \boldsymbol{C}_{s|h})$. Note that the multiple integral in Eq. (5.2) does not have the form of a multivariate Gaussian probability, as was the case in Eq. (4.7). As a matter of fact the numerical complexity of Eq. (5.2) may be much higher than that of Eq. (4.7) depending on the form of the additional term

$f_S(\boldsymbol{\chi}_{\text{soft}})$. This means that the trade-off for being able to process a wider class of uncertain knowledge is at the expense of a higher numerical cost.

5.3. The BME Mode Estimate

The BME mode estimate $\hat{\chi}_k$ is the most probable value for the S/TRF $X(\boldsymbol{p})$ at the estimation point $\boldsymbol{p}_k$ and it is obtained by maximizing the BME posterior pdf $f_K(\chi_k)$ as expressed by Eq. (4.8). Following a derivation similar that than of Chapter 4., Eqs. (4.9)-(4.11), we obtain the following equation for the BME mode estimate

$$\hat{\chi}_k = m_k + \frac{-1}{(\boldsymbol{C}^{-1}_{\text{map}})_{k,k}} \left(\sum_{i=1}^{m_h} (\boldsymbol{C}^{-1}_{\text{map}})_{i,k} (\chi_i - m_i) + \sum_{j=m_h+1}^{m} (\boldsymbol{C}^{-1}_{\text{map}})_{i,k} (\overline{\chi}_i - m_i) \right), \tag{5.3}$$

where $\overline{\chi}_i - m_i = \left(\int d\boldsymbol{\chi}_{\text{soft}}\ f_S(\boldsymbol{\chi}_{\text{soft}})\, f_G(\boldsymbol{\chi}_{\text{map}}) \right)^{-1} \int d\boldsymbol{\chi}_{\text{soft}}\ f_S(\boldsymbol{\chi}_{\text{soft}})(\chi_i - m_i)\, f_G(\boldsymbol{\chi}_{\text{map}})$ for i= $m_h + 1, m$.

When calculating $\overline{\chi}_i$ numerically, it is more efficient to use the following expression, where without loss of generality the assumption that m_i=0 has been made

$$\overline{\chi}_i = \left(\int d\boldsymbol{\chi}_{\text{soft}}\ f_S(\boldsymbol{\chi}_{\text{soft}})\, \phi(\boldsymbol{\chi}_s, \boldsymbol{m}_{s|kh} \boldsymbol{C}_{s|kh}) \right)^{-1} \int d\boldsymbol{\chi}_{\text{soft}}\ f_S(\boldsymbol{\chi}_{\text{soft}})\, \chi_i\, \phi(\boldsymbol{\chi}_s, \boldsymbol{m}_{s|kh} \boldsymbol{C}_{s|kh}), \tag{5.4}$$

where $\mathbf{m}_{s|kh} = \boldsymbol{C}_{s,kh} \boldsymbol{C}^{-1}_{kh,kh} \boldsymbol{\chi}_{\text{kh}}$ and $\mathbf{C}_{s|kh} = \boldsymbol{C}_{s,s} - \boldsymbol{C}_{s,kh} \boldsymbol{C}^{-1}_{kh,kh} \boldsymbol{C}_{kh,s}$. The numerical approach to calculate the BME mode estimate $\hat{\chi}_k$ is to assume an initial value $\hat{\chi}_k{}^0$, calculate a new $\hat{\chi}_k$ using the above Eqs. (5.3) and (5.4), and iterate until the solution is found. While the BME mode estimate $\hat{\chi}_k$ is the most probable value of $X(\boldsymbol{p})$ at the estimation point $\boldsymbol{p}_k$,

valuable information is also provided by calculating the statistical moments of the BME posterior pdf. This is considered next.

5.4. Moments of the BME Posterior PDF

As explained in the previous chapters the moments of $f_K(\chi_k)$ of interest are its mean $\bar{x}_{k|K}$, which is BME mean estimate minimizing error variance, its variance $\sigma^2_{k|K}$, also called error variance and is a measure of uncertainty associated with the BME mean estimate, and the third order moment which provides skewness information. Following along the lines of Chapter 4., Eqs. (4.12)-(4-21) we obtain similar equations for the mean, variance and skewness of the posterior, which are given as follow.

The mean $\bar{x}_{k|K} = \int \chi_k \chi_k f_K(\chi_k) = A^{-1} \int \chi_k \chi_k \int d\boldsymbol{\chi}_{\text{soft}}\, f_S(\boldsymbol{\chi}_{\text{soft}}) \phi(\boldsymbol{\chi}_{\text{map}}; \boldsymbol{0}, \boldsymbol{C}_{\text{map}})$ of the BME posterior pdf may be rewritten as

$$\bar{x}_{k|K} = A'^{-1} \int d\boldsymbol{\chi}_{\text{soft}}\, f_S(\boldsymbol{\chi}_{\text{soft}}) \boldsymbol{B}_{k|hs} \boldsymbol{\chi}_{\text{data}}\, \phi(\boldsymbol{\chi}_{\text{soft}}; \boldsymbol{B}_{s|h} \boldsymbol{\chi}_{\text{hard}}, \boldsymbol{C}_{s|h}) \tag{5.5}$$

where $\boldsymbol{B}_{k|hs} = \boldsymbol{C}_{k,hs} \boldsymbol{C}^{-1}_{hs,hs}$, $\boldsymbol{C}_{k|hs} = \boldsymbol{C}_{k,k} - \boldsymbol{B}_{k|hs} \boldsymbol{C}_{hs,k}$, $\boldsymbol{B}_{s|h} = \boldsymbol{C}_{s,h} \boldsymbol{C}^{-1}_{h,h}$, $\boldsymbol{C}_{s|h} = \boldsymbol{C}_{s,s} - \boldsymbol{B}_{s|h} \boldsymbol{C}_{h,s}$ and $A' = \int d\boldsymbol{\chi}_{\text{soft}}\, f_S(\boldsymbol{\chi}_{\text{soft}}) \phi(\boldsymbol{\chi}_{\text{soft}}; \boldsymbol{B}_{s|h}\, \boldsymbol{\chi}_{\text{hard}}, \boldsymbol{C}_{s|h})$. Eq. (5.5) can also be expressed in the more compact form of

$$\bar{x}_{k|K} = \mathbf{B}_{k|hs(h)} \boldsymbol{\chi}_{\text{hard}} + \mathbf{B}_{k|hs(s)} \bar{\boldsymbol{\chi}}_{\text{soft}}, \tag{5.6}$$

where $\bar{\boldsymbol{\chi}}_{\text{soft}} = A'^{-1} \int d\boldsymbol{\chi}_{\text{soft}}\, \boldsymbol{\chi}_{\text{soft}}\, f_S(\boldsymbol{\chi}_{\text{soft}}) \phi(\boldsymbol{\chi}_{\text{soft}}; \boldsymbol{B}_{s|h} \boldsymbol{\chi}_{\text{hard}}, \boldsymbol{C}_{s|h})$, which, in the limiting case where only hard data is used, simplifies to the Simple Kriging estimator, Eq. (4.15).

Similarly the variance of the BME posterior pdf is given by

$$\sigma^2_{k|K} = \boldsymbol{C}_{k|hs} + A'^{-1} \int d\boldsymbol{\chi}_{\text{soft}}\, f_S(\boldsymbol{\chi}_{\text{soft}})(\mathbf{B}_{k|hs}\boldsymbol{\chi}_{\text{data}} - \bar{x}_{k|K})^2 \phi(\boldsymbol{\chi}_{\text{soft}}, \mathbf{B}_{s|h}\boldsymbol{\chi}_{\text{hard}}, \boldsymbol{C}_{s|h}), \qquad (5.7)$$

and order moment $\mu_{k,3|K}$ of the posterior pdf is given by

$$\mu_{k,3|d} = A'^{-1} \int d\boldsymbol{\chi}_{\text{soft}}\, f_S(\boldsymbol{\chi}_{\text{soft}})(\mathbf{B}_{k|hs}\boldsymbol{\chi}_{\text{data}} - \bar{x}_{k|K})^3 \phi(\boldsymbol{\chi}_{\text{soft}}, \mathbf{B}_{s|h}\boldsymbol{\chi}_{\text{hard}}, \boldsymbol{C}_{s|h}). \qquad (5.8)$$

5.5. Numerical Implementation

The integrand of the multiple integral in Eq. (5.1) is much more challenging than that of Eq. (4.1) corresponding to interval soft data , because $f_S(\boldsymbol{\chi}_{\text{soft}})$ can take any form the user specifies. It is possible to use a general-purpose algorithm for numerical approximation of multiple integrals which would work for any form of $f_S(\boldsymbol{\chi}_{\text{soft}})$, however such algorithm are numerically complex, and are usually computationally expensive (Berntsen, *et al.*; 1991). Another approach is to use different numerical implementations for different forms of the soft pdf $f_S(\boldsymbol{\chi}_{\text{soft}})$. This approach allows to simplify the numerical complexity of Eq. (5.1) by taking advantage of the specific form of the soft pdf. Consider the following example.

EXAMPLE 5.2: Assume that the soft pdf $f_S(\boldsymbol{\chi}_{\text{soft}})$ may be written as a product of independent univariate pdf's for each of the soft data points, i.e. $f_S(\boldsymbol{\chi}_{\text{soft}}) = \prod_{i=m_h}^{m} f_{S,i}(\chi_i)$, and that each $f_{S,i}(\chi_i)$ may be expressed by a step-wise function, i.e. $f_{S,i}(\chi_i) = \sum_{j=1}^{N_i} f_{i,j}\delta_{l_{i.j} \le \chi_i < l_{i.j+1}}$. Then the posterior pdf of Eq. (5.1) may be expressed as $f_K(\chi_k) = A^{-1} \prod_{i=m_h}^{m} \sum_{j=1}^{N_i} f_{i,j} \int_{l_{i,j}}^{l_{i,j}} d\chi_i f_G(\boldsymbol{\chi}_{\text{map}})$ which is the sum of multiple integrals of the form of Eq. (4.1). As explained in the previous chapter, each multiple integrals in this sum is known as a multiple Gaussian probability, and may be efficiently calculated using the same

numerical libraries as for BMEintEst. Provided that the number of soft data points is small, this approach will be more efficient than using a general purpose algorithm for multiple integrals.

The BME equations for spatiotemporal estimation presented in this chapter were implemented numerically for several synthetic test cases and case studies. For each test case and case studies a different numerical implementation was developed that took advantage of the specific form of the soft pdf considered. These implementations are each specific to a particular test case or case study, but globally they will be referred to as the BMEpdfEst code, so as to distinguish from the BMEintEst code of last Chapter.

By way of a summary, the input to the programs of the `BMEpdfEst code` includes the mean and covariance function, as well as hard and soft data of pdf type. The output of the programs include, (i) the mapping estimates (mode and mean of the posterior pdf), (ii) the variance of the posterior pdf (simple assessment of mapping accuracy), (iii) the BME confidence intervals (single-point error assessment).

The most efficient numerical library depends on the form of the soft pdf $f_S(\boldsymbol{\chi}_{\text{soft}})$. If the soft pdf has a shape that is not complex, (such as in Example 5.2) and the number of soft data points is sufficiently small, then using the same numerical library as for interval soft data may be more efficient. However if the shape of the soft pdf is complicated, a general purpose method to approximate multidimentional integrals may be better. As a result the numerical work may vary wildly from case to case, with a lower bound corresponding to soft data of the interval type.

5.6. Simulated Case Studies

Several synthetic case studies were conducted to compare the BME method with traditional kriging methods in the context of soft data of probabilistic type. The Indicator Kriging was found to perform poorly compared to BME method when using soft data of the interval type. The Simple Kriging method (SK) and the Simple Kriging with Measurement Error (SKME) methods were found to perform better in the previous Chapter, so they are the method used here to compare with BME when using soft data of the probabilistic types. Three synthetic test cases are presented in the following. The first comparison is based on stochastic simulation which allows to compare the accuracy of estimated values between the Simple Kriging methods and BME. The two subsequent case studies are oriented toward mapping applications.

5.6.1. A Simulation Based Comparison Between BME and Kriging Methods

Stochastic simulations of a S/TRF are useful constructs allowing to compare the accuracy of different estimation methods. Detailed explanations of stochastic simulation methods and the SK and SKME methods are given in Chapter 2., however for the sake of self containment these methods are summarize first in the current context before giving the results of the comparison.

Simulation of S/TRF with pdf soft information

In order to test and compare different spatiotemporal estimation methods, it is useful to generate several realization of the S/TRF with hard data and soft data of the probabilistic type. Each realization of the field consists of a value $\boldsymbol{\chi}_{\text{hard}}^{(\ell)}$ for the hard data vector, and the value of the soft pdf $f_S^{(\ell)}(\boldsymbol{\chi}_{\text{soft}})$. where the subscript ℓ denotes a specific realization of the field, $\ell = 1,...,L$. Using the technique described in Table 2.2 it is possible to generate

realizations of hard data $\boldsymbol{\chi}_{\text{hard}}^{(\ell)}$ and soft probabilistic data $f_S^{(\ell)}(\boldsymbol{\chi}_{\text{soft}})$ by specifying a *generator pdf* $f_v(\upsilon)$. The generator pdf $f_v(\upsilon)$ represents the uncertainty associated with the measurement value at the soft data point, and the soft pdf $f_{S,i}^{(\ell)}(\chi_{s,i})$ for each soft data point is obtained by translation of $f_v(\upsilon)$ as shown in Fig. 2.4.

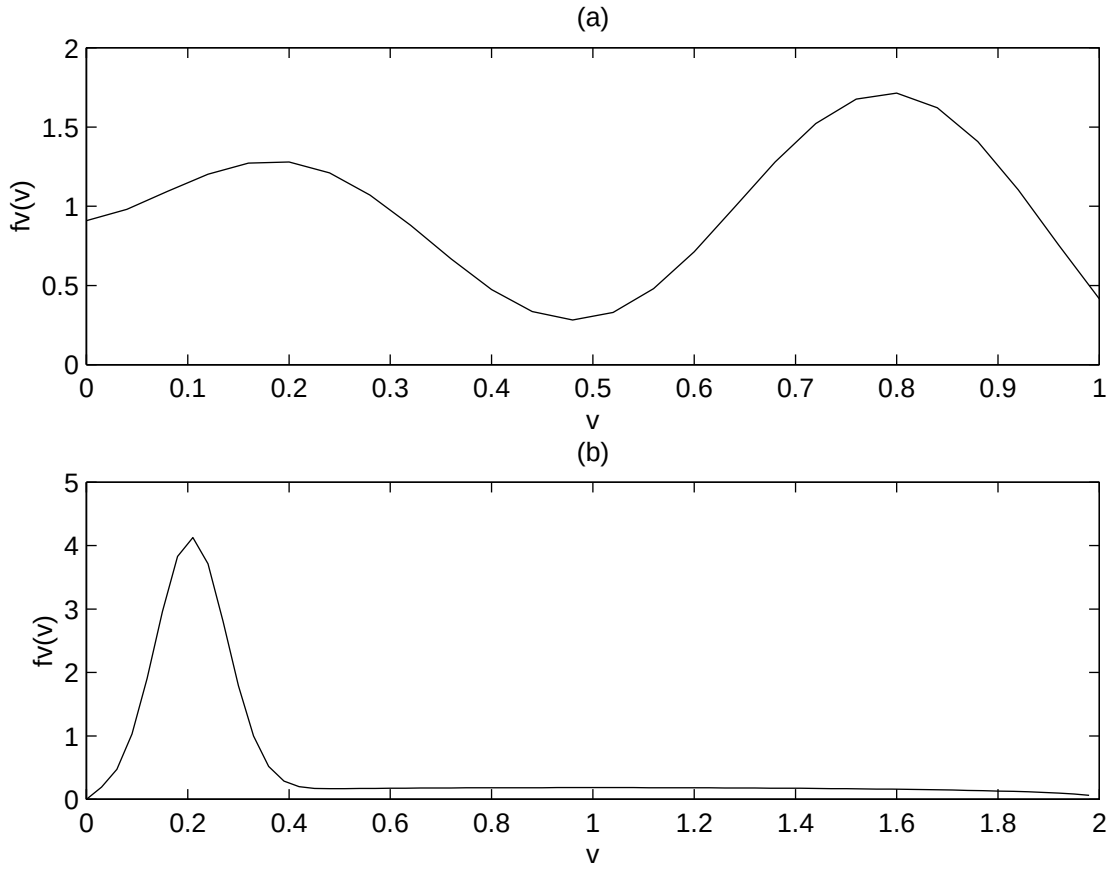

Figure 5.1: Probability distribution functions $f_v(\vartheta)$ used to generate the soft pdf realizations for: (a) case 1 and (b) case 2.

Two cases of soft generator pdf $f_v(\upsilon)$ are investigate in this case study. The shape of $f_v(\upsilon)$ for these two cases is shown in Fig. 5.1. For each of these soft generator pdf several realizations of the hard and soft data were generated for a S/TRF with a zero mean and the covariance model $c_x(r) = c_o \exp[-r^2 / a_r^2]$ where a_r=1.0 and c_o=1.0. The location of the 2 hard data points and the 2 soft data points used in each realization are shown in Fig. 5.2.

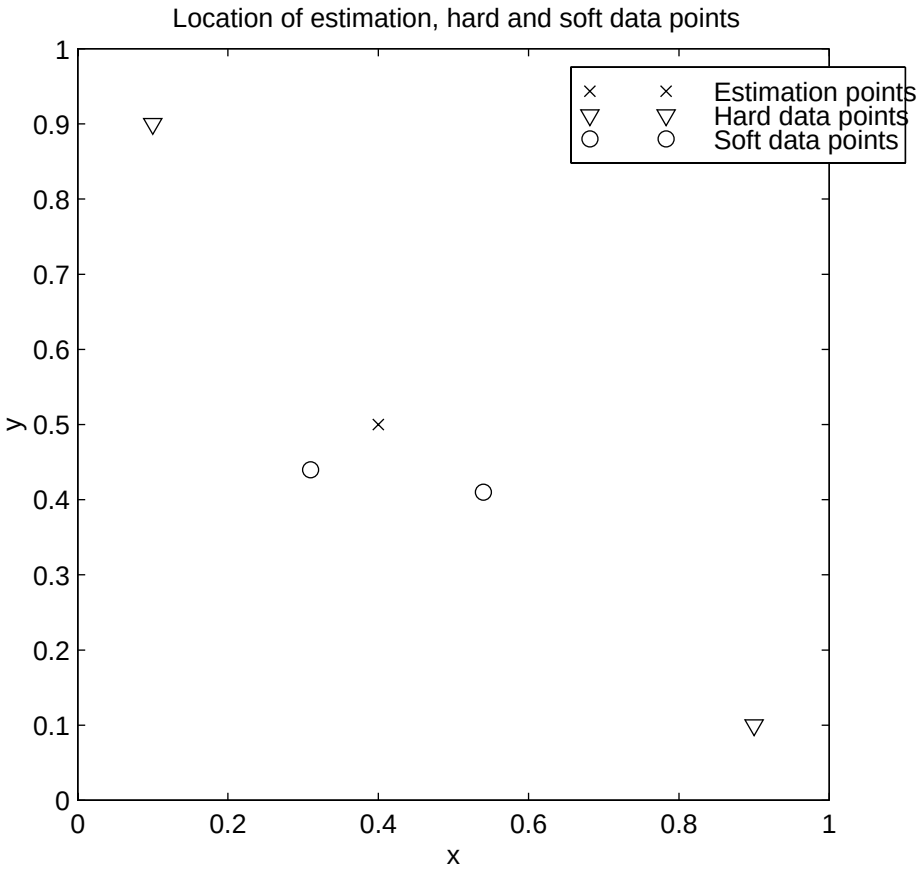

Figure 5.2: Location of the hard data points, soft data points and estimation point.

Prediction using the simple kriging methods

The first method considered, called SKh, only takes in account the hard data. As showed in Chapter 2., Eqs. (2.28)-(2.34) the estimator for SKh is $\chi^*_{k,SKh} = \boldsymbol{C}_{k,h}\,\boldsymbol{C}^{-1}_{h,h}\,\boldsymbol{\chi}_{\text{hard}}$, and the error variance is $\sigma^2_{e,SKh} = C_{k,k} - \boldsymbol{C}_{k,h}\boldsymbol{C}^{-1}_{h,h}\boldsymbol{C}_{k,h}$

The second method considered, SKME, accounts somewhat approximately for soft data. As explained in Chapter, Eqs. (2.40)-(2.46) the estimator for SKME is $\chi^*_{k,SKME} = \boldsymbol{C}_{k,hs}\,\boldsymbol{C}^{-1}_{hs,hs(\sigma_V)}[\boldsymbol{\chi}_{\text{hard}}, \overline{\boldsymbol{\chi}}_{\text{soft}}]$ where the vector $\overline{\boldsymbol{\chi}}_{\text{soft}}$ has components equal to $\overline{\chi}_{\text{i}} = \int d\chi_{\text{i}}\,\chi_{\text{i}}\,f_{S,i}(\chi_{\text{i}}), \quad (i = m_h + 1,...,m)$, and $\boldsymbol{C}^{-1}_{hs,hs(\sigma_V)}$ is obtained by adding $\sigma^2_V = \int d\chi_{\text{i}}(\chi_{\text{i}} - Y_i)^2\,f_{S,i}(\chi_{\text{i}})$ in the diagonal elements of the matrix $\mathbf{C}_{hs,hs}$ corresponding

to the soft data point i, $(i = m_h + 1, \ldots, m)$. The error variance for the SKME method is given by $\sigma^2_{e,SKh} = C_{k,k} - \boldsymbol{C}_{k,hs}\boldsymbol{C}^{-1}_{hs,hs(\sigma_V)}\boldsymbol{C}_{k,hs}$.

The confidence interval for a of given confidence probability η is estimated by assuming that the estimated variable is gaussian with mean χ^*_k (estimator of the method), and variance σ^2_e (error variance of the estimation method). Thus, the η-confidence interval is given by $[\Phi^{-1}((1-\eta)/2; \chi^*_k, \sigma^2_e), \quad \Phi^{-1}(1-(1-\eta)/2; \chi^*_k, \sigma^2_e)]$, where $\Phi^{-1}(x; \bar{x}, \sigma^2)$ is the inverse of the gaussian cdf with mean $\bar{x}$ and variance σ^2.

Results of the numerical comparison between BME and Simple Kriging methods

In this synthetic case study we compare the SK and SKME estimators with the BME mode estimator (Eq. 5.3), called BMEpdfMode, and the BME mean estimator (Eq. 5.5), called BMEpdfMean. In order to compare the methods using stochastic simulation we generate several realizations of a S/TRF, each realization includes the hard data, soft data, and the value of the S/TRF at the estimation point. The value at the estimation point is interpreted as the "true" value, and it is ignored in the estimation process. Using only the hard data and soft data this value is "re-estimated", and the difference between re-estimated value and true value represents the estimation error. The accuracy of an estimation method is assessed by looking at the average and distribution of the estimation errors: The better an estimation error, the smaller is its estimation error.

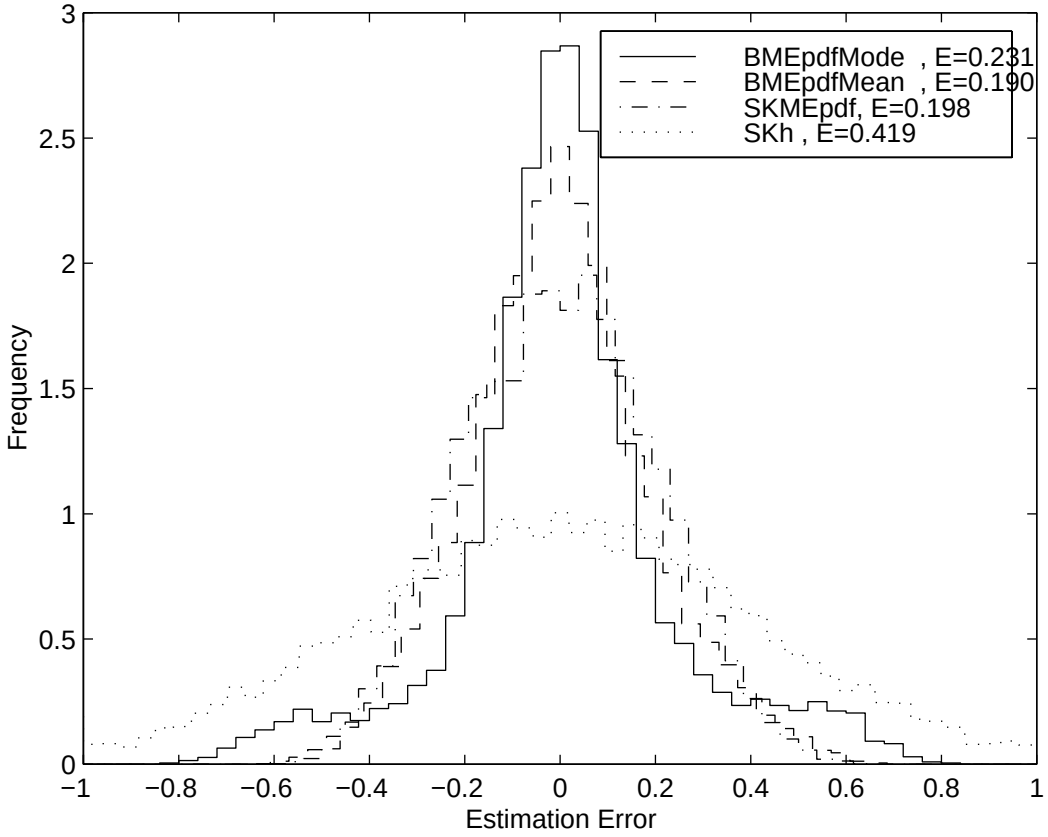

Figure 5.3: Histogram of prediction errors for the BMEpdfMode, BMEpdfMean, SKME and SKh, case 1

The distributions of estimation errors obtained using 10,000 realizations of the S/TRF obtained with case 1 of the generating pdf $f_v(\vartheta)$ (Fig 5.1.a) are shown in Fig 5.3 for the BMEpdfMode, BMEpdfMean, SKME and SKh methods. Also shown in the legend of Fig 5.3 is the average estimation error E for each of the method. As one can see the BME estimators perform better than both kriging methods. The average error of the BME estimators (E=0.231 for BMEpdfMode and E=0.190 for BMEpdfMean), are about half of the average error E=0.419 for the SKh method. As expected, the SKME method performs better than the SKh method. However it is clear from the figure that the distribution of estimation errors for the BME estimators have a sharper peaked shape around zero than that of the SKME estimator. This means that the BME estimators are

more likely to a give correct estimation than the SKME estimator, and hence they are better estimators with respect to that criteria. It is interesting to note that the BME mode estimator has a higher peak while the BME mean estimator has a smaller average estimation error E. This is expected since the mode estimator is the most likely estimation while the (conditional) mean estimator minimizes the error variance. As a result BME offers the flexibility to choose the estimator that is best suited for the mapping situation. On the other hand kriging methods such as the SKME method do not offer this flexibility: They are estimator which minimizes the error variance. Hence if we want to compare SKME and BME according to the average estimation error criteria, we should compare the SKME estimator and the BME mean estimator. Doing so revels that, again, with a smaller average estimation error of E=0.190, BMEpdfMean is a better estimator than SKME. We conclude from this test case that the BME method are better than the SKME and SK methods. We also note that SKME seems to capture a large part of the soft information and hence works substantially better than SKh. This is actually due to the shape of the soft pdf. Compare the soft pdfs of Fig 5.1(a) and Fig 5.1(b). The soft pdf of case 1 is less informative than in case 2 because it merely states that the attribute has a value between 0 and 1, while in case 2 the sharper peak give a narrower range of likely values for the attribute. Consequently case 1 represents an "easy" mapping situation where an approximate estimator such SKME might provide an acceptable estimation (because the soft information is not informative), while case 2 is a more "difficult" mapping situation, where even the SKME method is going perform poorly, as shown next.

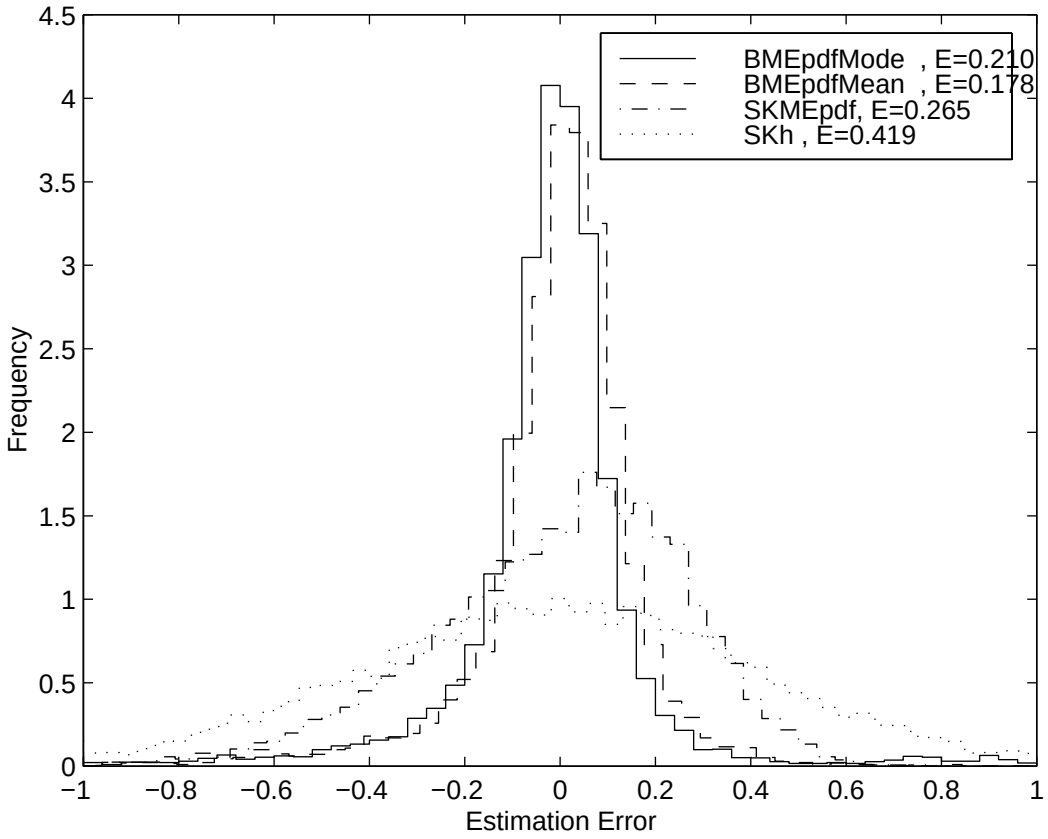

Figure 5.4: Histogram of prediction errors for the BMEpdfMode, BMEpdfMean, SKME and SKh, case 2

The distributions of estimation errors obtained with case 2 of the generating pdf $f_v(\vartheta)$ (Fig 5.1.b) are shown in Fig 5.4 for the BMEpdfMode, BMEpdfMean, SKME and SKh methods. Also shown in the legend of Fig 5.4 is the average estimation error E for each of the method. It is evident from the figure that in this case the BME estimators are much better than either the SK and SKME methods. In terms of the distribution of estimation errors, the peaks of the curves for the BME estimators are more than twice as high than either SK or SKME. This means that the BME estimators are more than twice as likely to give a correct estimation than the SK and SKME methods. Even more interesting is the fact that SKME does not seem to be performing much better than SK according to that criteria. When using the average estimation error as comparison criteria, we find that

the average error of E=0.265 for SKME for SKME represents a substantial increase over the E=0.178 for the BME mean estimate. This results show that when using a combination of hard data and soft data of the probabilistic type, the BME method produces estimates that are consistently better than any existing kriging methods, and may be substantially better depending on how informative is the soft data.

A more informative assessment of estimation accuracy is provided by the confidence intervals. As explained earlier the confidence intervals for the SK and SKME methods are obtained from the estimated error variance by assuming a Gaussian distribution for the estimated random variable. These intervals are centered around the estimated values (i.e., $\chi^*_{k,\mathrm{SK}}$ and $\chi^*_{k,\mathrm{SKME}}$) and have a length directly calculated from the estimation error variance (i.e., σ^2_{SK} and σ^2_{SKME}). BME confidence intervals, on the other hand, are directly derived from the BME posterior pdf using high probability density sets, as given by Eq. (3.37). Typical BME confidence intervals obtained using Eq. (3.37) are shown in Fig. 3.2. On the basis of the approach described above, we calculated the confidence intervals for 1,000 realizations, using a confidence probability of η=0.1, 0.5, and 0.9 for the BME, SK, and SKME methods, respectively. The calculated confidence intervals honored the data. In the case, e.g., of the 90% confidence intervals, we found that the SK and SKME estimated values were within their respective confidence intervals for 90% of the realizations, while the BME mode estimator was within the 90% BME confidence interval for 90% of the realizations, as well. Next, we compared the length of the calculated intervals. In Table 5.1 we show the average length of the BME confidence intervals and of the confidence intervals calculated using SKME. It is interesting to note that the lengths of the BME confidence intervals are much smaller than these obtained by the SKME methods (by as much as 50%). This illustrates exactly the claim made earlier that, BME confidence intervals are accurate and yet substantially smaller than the confidence intervals calculated using traditional methods. This aspect is critical in several

mapping situations where higher costs are associated with larger confidence intervals (e.g., contamination mapping of hazardous waste sites where cleaning up over a wider range of contamination levels considerably increases the remediation cost).

TABLE 5.1: Average length of confidence intervals

Confidence Probability	Length of BME confidence intervals	Length of confidence intervals for SKME
0.1	0.0354	0.0736
0.5	0.1791	0.3951
0.9	0.4320	0.9634

5.6.2. Map Based Comparison Between BME and Kriging Methods

In this problem set we compare the mean estimator of the BME method (referred from now on as the BME estimator) with the SK and SKME estimators by re-estimating known values of a Spatial Random Field (SRF). The known (true) values are obtained by simulating the values of the SRF at the nodes of a regular grid. A set of hard and soft data points is selected at random from the grid nodes to form the knowledge base for the estimation methods. Using only the hard and soft data points the values at the other grid nodes are re-estimated for each of the estimation methods. Estimation errors are obtained by subtracting the known values from the estimated values. A comparison of estimation errors shows that the BME method is providing a better prediction than the SK and SKME methods, and that the prediction of estimation errors provided by the BME method is more informative than that provided by the kriging methods. Following we explain how the data for this case were generated, and we present the numerical results obtained for two different realizations of the SRF.

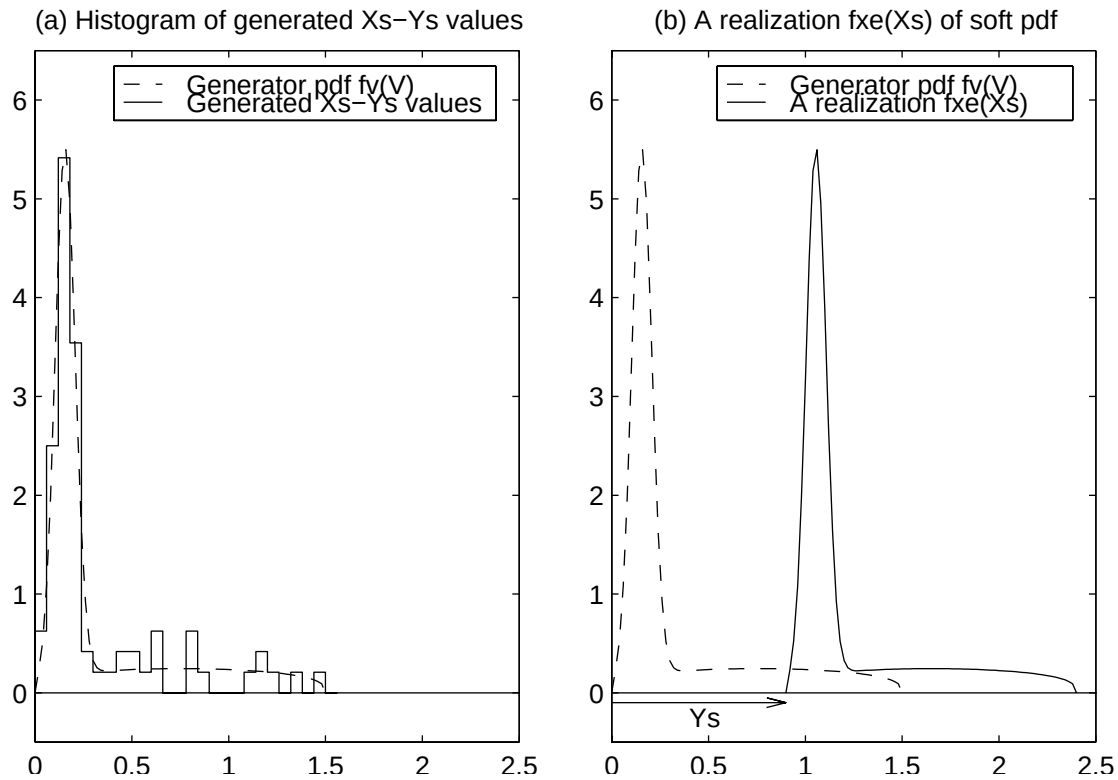

Figure 5.5: (a) Histogram of values $\vartheta_i=(\chi_{s,i}-\psi_{s,i})$, i=1,...,80 generated using the generator pdf $f_v(\vartheta)$, (b) a realization of the soft pdf $f_{x(e)}(\chi_{s,i})$ obtained for a given Ys.

The generator pdf $f_v(\vartheta)$ used in this case study for the stochastic simulation method described in 5.6.1. is shown with dashed line in Fig. 5.5a. The stochastic simulation method starts by generating the values of the SRF are on the 21x21 nodes of the grid shown in Fig 5.6a for a SRF with zero mean and covariance model $c_x(r)=c_n+c_o\exp[-r^2/a_r^2]$, where c_n=0.02, c_o=0.98 and a_r=1.0. A set of 8 hard data points and 80 soft data points were randomly selected as shown in Fig 5.6a with triangles and circles, respectively. Using these points, we obtain realizations of hard data values $\boldsymbol{\chi}_{\text{hard}}^{(\ell)}$ and of the soft pdfs $f_S^{(\ell)}(\boldsymbol{\chi}_{\text{soft}})$, $\ell=1,\ldots,L$, where ℓ denotes a realization. The hard data points are assumed to represent measurements obtained with an accurate but expensive device, while the measurement at the soft data points are assumed to come from an inaccurate but supposedly inexpensive source. This mapping situation may represent e.g., the case of nitrate concentration in an aquifer coming from core samples (expensive and

accurate), and from "quick and dirty" measurements at location where the ground water is believed to seep to the ground surface, i.e. springs or ponds (inexpensive but inaccurate measurement). The larger number of soft data available is representative of the lower cost of such information, which is realistic in the context of the trade-off between accuracy and cost for most measurement campaign.

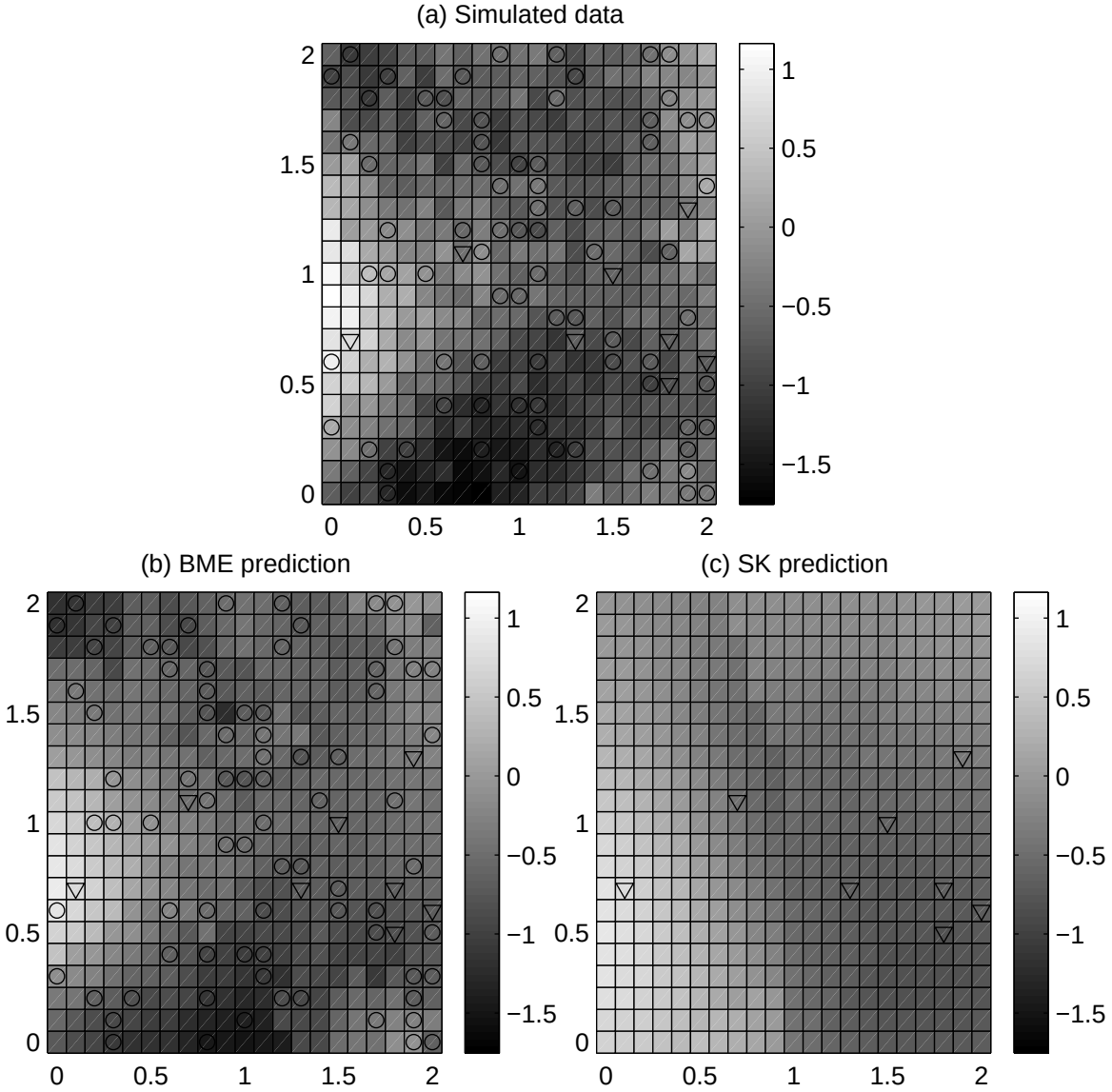

Figure 5.6: (a) Realization of the simulated SRF. The hard and soft data point used for the estimation are shown with triangles and circles respectively. (b) BME prediction and (c) SK estimation.

A typical realization of the SRF obtained with the simulation technique described above is shown in figure 5.5a. Using only the hard data values and soft information, values at the grid nodes are re-estimated using the BME, SK and SKME estimation methods. When doing the estimation, the BME and SKME methods use the 2 hard and 2 soft data points closest to the estimation point, while the SK uses only the 2 hard data points closest to the estimation point. The estimated values obtained with the BME and SK methods are shown in figure 5.6b and 5.6c respectively, and the histograms of estimation errors for the BME, SKME and SK methods are shown in figure 5.7a. As one can see from figure 5.6 the BME prediction map reproduces well the true map whereas the SK prediction map is a poor estimation of the true data. The square root of the average mean square error E, shown in the legend of figure 5.7a, is lower for the BME prediction, while it is 25% and 128% larger for SKME and SK respectively.

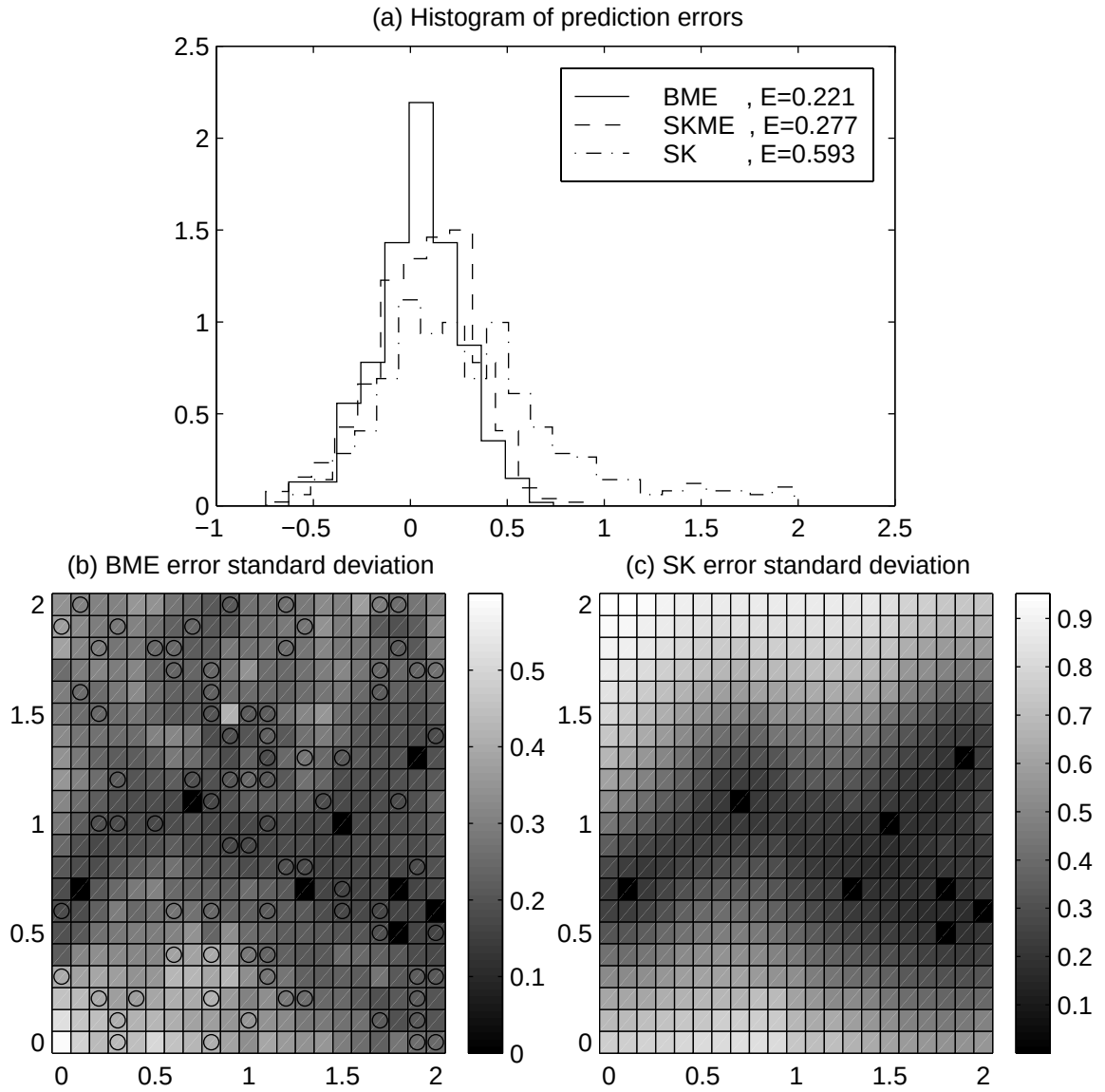

Figure 5.7: (a) Histogram of prediction errors for BME, SKME and SK showing the Average mean square error E, (b) predicted BME error standard deviation, and (c) predicted SK error standard deviation.

The estimation methods also provide an estimate of the estimation errors. In figure 5.7b and 5.7c we show the predicted error standard deviation for the BME and SK methods. By comparing the scales one can see that the predicted error standard deviation is about twice as large for SK than for BME.

A second realization of the SRF is used to compare the estimation methods as shown in figures 5.8 and 5.9. Comparing the simulated (true) data in Fig. 5.8(a) with the BME and SK predictions in figure 5.8b and 5.8c we again see that BME provides a much better prediction than SK. The histogram of prediction errors in Fig. 5.9a shows that BME is again more accurate than both SKME and SK. Finally one interesting aspect of the error standard deviation shown in Fig. 5.9b and 5.9c is that the BME error standard deviation for this realization is different from that of the previous realization shown in figure 5.7b. The ability for the BME method to predict an error estimate that is a function of the data is one of the strength of the BME method, unlike the error estimates of both SKME and SK which are only of function of the hard and soft point locations.

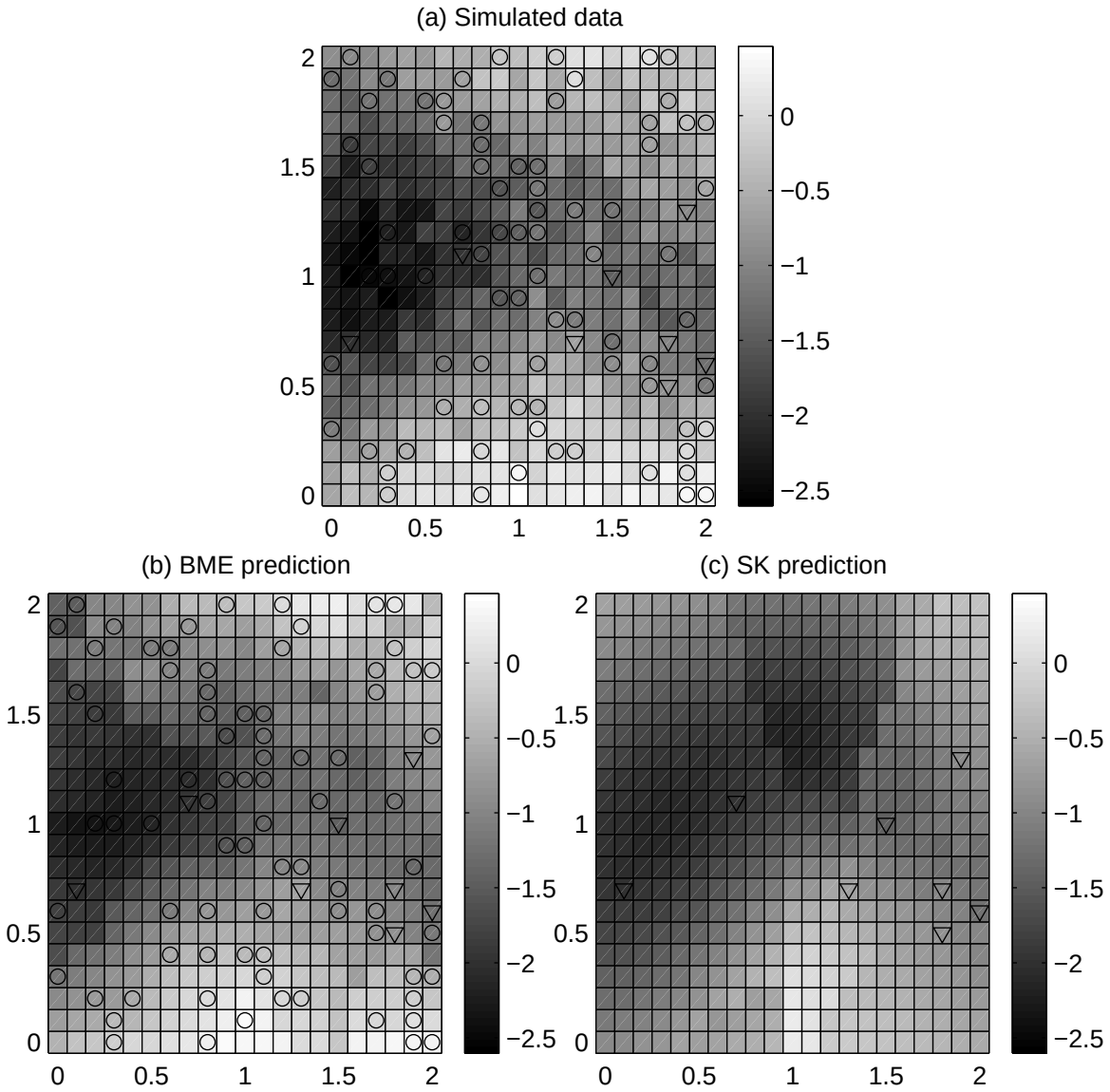

Figure 5.8: (a) Realization of the simulated SRF. The hard and soft data point used for the prediction are shown with triangles and circles respectively. (b) BME prediction and (c) SK prediction.

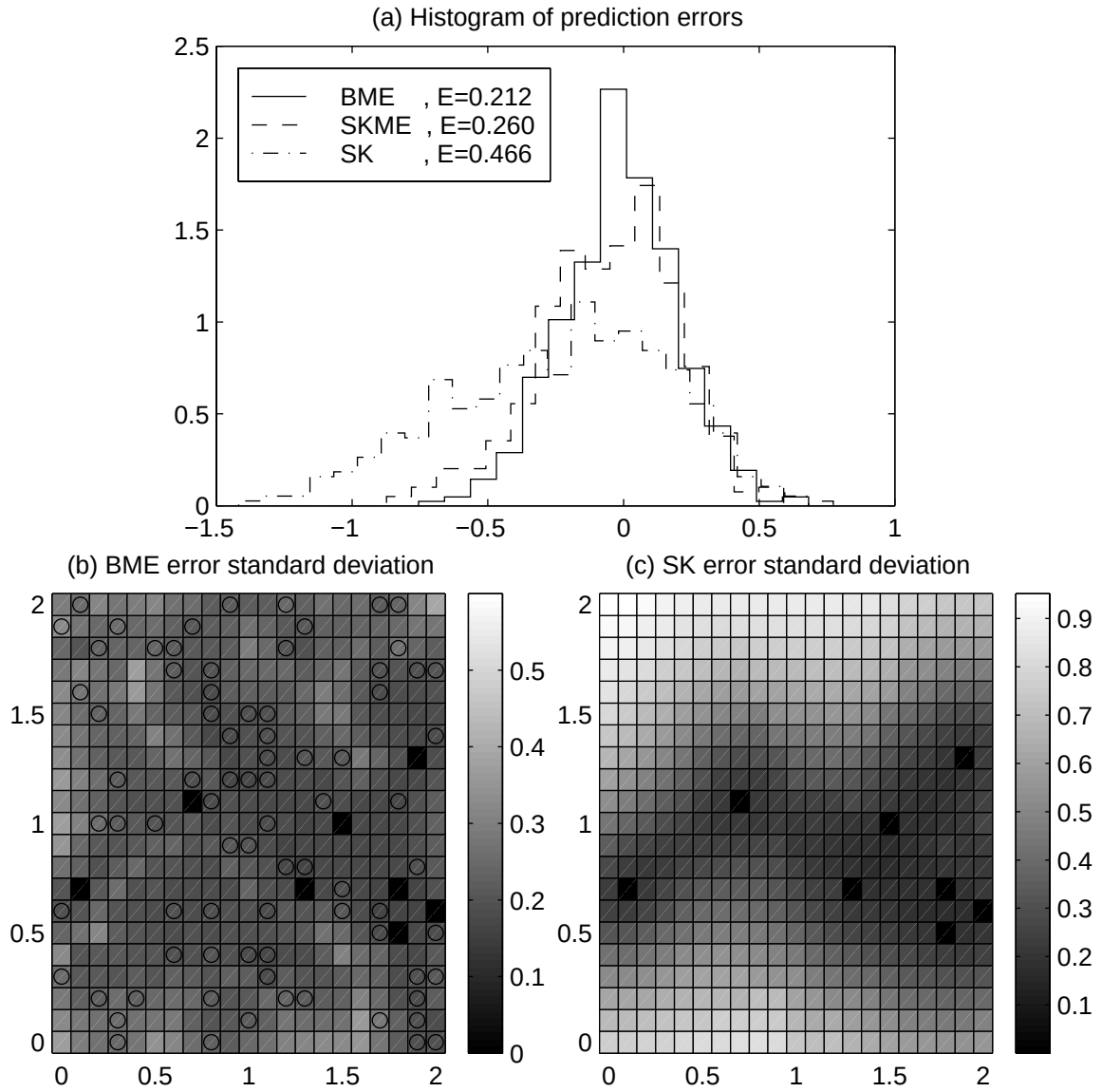

Figure 5.9: (a) Histogram of prediction errors for BME, SKME and SK showing the Average mean square error E, (b) predicted BME error standard deviation, and (c) predicted SK error standard deviation.

5.7 The Equus Beds Case Study

The Equus-Beds aquifer is an alluvial deposit near the city of Wichita in South-Central Kansas. The Wichita well-field was installed in the Equus-Beds aquifer to supply water to the city and pumping started in 1940. Ground water pumping from the well-field and the droughts during the 1950's and late 1980's resulted in a substantial decline of water-levels over a large area. The decline of this vital resource has motivated regulatory agencies to monitor the water-levels using a network of ground water observation wells (Olea, 1982). Measurements of the water-level at the observation wells were made throughout the years. However, because of recording errors and due to the difficulties of measuring accurately a fluctuating water-level in a pumping well-field, the information available consists of a combination of hard and soft (uncertain) data. The purpose of this case study is to incorporate both hard and soft data into BME analysis in order to produce accurate water-level spatiotemporal maps and perform reliable error assessment. These maps can improve the hydrogeologic understanding of the entire region and optimize the local decision making regarding the operation of the Wichita well-field.

The study area (Fig. 5.10) covers approximately 4000 Km^2 (latitudes $37°44'$ and $38°27'$; longitudes $-98°08'$ and $-97°22'$). The latitude and longitude coordinates were converted to Northing and Easting (in Km) using a projection procedure where the origin was set at $40°$ latitude and $-100°$ longitude. A total of 70 observation wells were used; these wells were numbered from 1 to 70 by increasing Northing (see, Fig. 5.10). The Little Arkansas river flows from North-West to South-East through the study area, and the land surface slopes towards the river, as is depicted by the contour lines of equal land surface elevation above the sea level (in ft; Fig. 5.10).

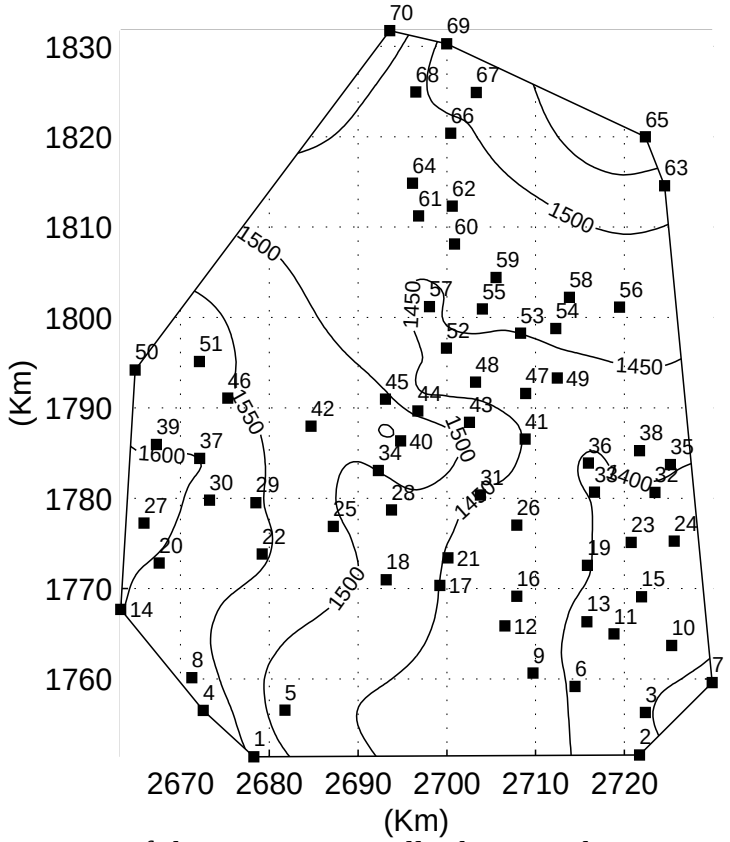

Figure 5.10: Locations of the monitoring wells shown with squares and well numbers; and contour lines for the ground elevation annotated in *ft*. The Northing (vertical axis) and Easting (horizontal axis) are in *Km*

The water-level decline in the Equus Beds aquifer has been documented in previous works. Much of this decline occurred from 1940 to early 1957 (Stramel, 1966). Water-levels stabilized in the 1960's and 1970's, continued to decline between the late 1970's and the 1988-92 drought, and reached their maximum decline to date of as much as 40 *ft* or more during 1991-1993. The water-level recovered moderately during the time period 1993-1998, primarily as a result of decreased city withdrawals (Aucott and Myers, 1998). Since 1995 the city of Wichita has investigated the possibility of artificial ground water recharge in the well-field to meet future needs and to protect the aquifer from saltwater intrusion from natural and anthropogenic sources to the West.

Water-level data have been collected by the city of Wichita personnel at the 70 wells mentioned above. The data collection started in 1940 and, as the well-field development proceeded, water-levels in additional wells were measured. Data were stored by the city in paper and electronic form and by the USGS in electronic form. Measurements are taken during the winter, when irrigation stops for a few months in order to minimize the draw down cone effects due to pumping in the well itself or at surrounding wells (Olea 1982). Measuring frequency varies from well to well, resulting in duplicate measurements for some winters and no measurements during other winters. In this work we used a dataset of 1,573 water-level elevation measurements provided by the Kansas Geological Survey (KGS). The dataset covers the period 1970-1998 for the observation wells shown in Fig. 5.10.

One aspect of particular importance to the KGS is that of the quality of measurements. Measurements of water-level elevations at each observation well are obtained by measuring the depth to water relative to a fixed measuring point, which is usually located a few *ft* above the land surface. Experience has shown that some measurements contain random errors due to several factors, including inaccurate readings, recording errors, uncertainty about the measuring point used for early measurements, or fluctuation of the water-level elevation during the monitoring season. Observations that did not have any attached measurement error were considered hard data. However, observations that had associated uncertainties were treated as soft data of either soft or probabilistic type by the BME approach. Soft data of the interval and probability types were assigned based on a review of the available readings and in light of the experience accumulated by the KGS in collecting similar data at other sites. Consider, e.g., the time profile of well no. 19 (Fig. 5.11; actual measurements are shown with circles). In this case, the duplicate measurements taken at the same monitoring season show fluctuations that are attributed to inaccurate readings and pumping-induced fluctuations. Limits on the

measurements are assigned, resulting in soft data of interval and probabilistic types, represented respectively by error bars and pulse-shaped curves in Fig. 5.11.

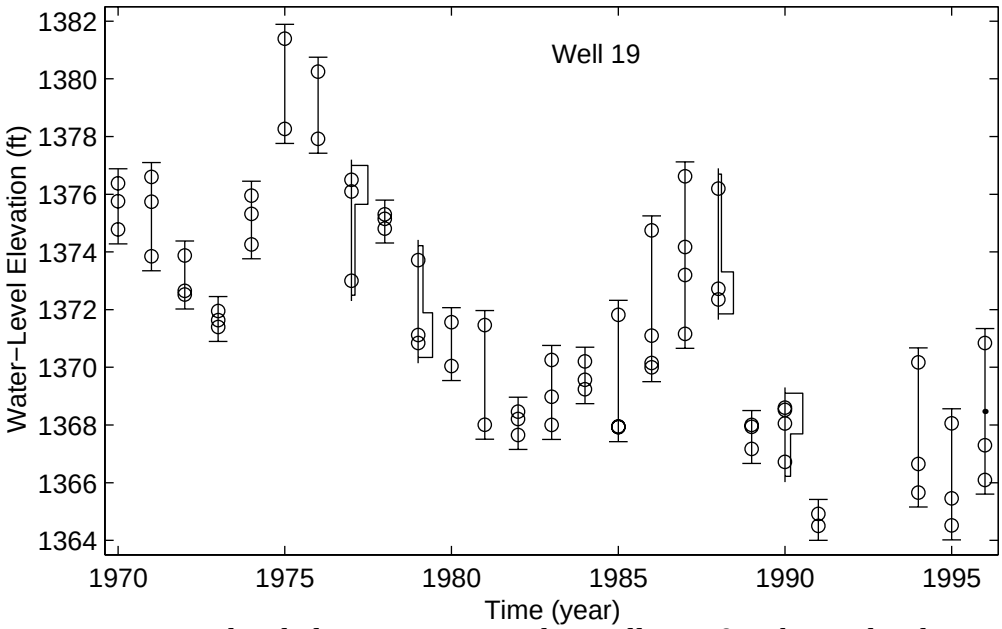

Figure 5.11: Water-level elevation measured at well no. 19. The circles depict available readings, the error bars show the interval (soft) data, and the pulse-shaped curves represent the probabilistic (soft) data.

The water-level elevation is modelled as an S/TRF $W(\boldsymbol{p})$, with space/time coordinates $\boldsymbol{p} = (\boldsymbol{s}, t)$; the spatial coordinates $\boldsymbol{s} = (s_1, s_2)$ specify Northing and Easting (in Km), and t is the temporal coordinate (in years). Physical considerations about water-level elevations suggest that the $W(\boldsymbol{p})$ is adequately represented as

$$W(\boldsymbol{p}) = \mu(\boldsymbol{s}) + X(\boldsymbol{p}), \tag{5.9}$$

where $X(\boldsymbol{p})$ is a homogeneous S/TRF with zero mean and separable covariance, and the mean value (or drift) of the water-level elevation $\mu(\boldsymbol{s}) = \overline{W(\boldsymbol{p})}$ is a function of the spatial location $\boldsymbol{s}$ only. The $\mu(\boldsymbol{s})$ strongly depends on the land surface elevation, and it is estimated at each well by taking the mean value of water-level elevation for that well (The

average number of observations used to calculate the mean water-level elevation at each well was greater than 35). The $X(\boldsymbol{p})$ contains all the randomness associated with the water-level elevation. In view of the homogeneity characteristics of $X(\boldsymbol{p})$, its covariance is estimated using the values obtained by subtracting the mean $\mu(\boldsymbol{s})$ from the observations of the water-level elevation $W(\boldsymbol{p})$. The covariance of $X(\boldsymbol{p})$ at the Wichita well-field does not have a strong anisotropy component, therefore an isotropic model was assumed. The covariance $c_x(r,\tau)$, $r = |\boldsymbol{s} - \boldsymbol{s}'|$ is the spatial lag and $\tau = |t - t'|$ the time lag, is estimated by

$$c_x(r,\tau) \approx \frac{1}{N(r,\tau)} \sum_{i=1}^{N(r,\tau)} (X_{-(r,\tau),i}\, X_{+(r,\tau),i}) - m_{-(r,\tau)} m_{+(r,\tau)}, \tag{5.10}$$

where $N(r,\tau)$ is the number of pairs of observed values $(X_{-(r,\tau),i}, X_{+(r,\tau),i})$ separated by the spatial and temporal lags r and t, $m_{-(r,\tau)}$ is the mean of the $X_{-(r,\tau),i}$ values and $m_{+(r,\tau)}$ is the mean of the $X_{+(r,\tau),i}$ values. The estimated values of $c_x(r,\tau)$ are shown in Fig. 5.12

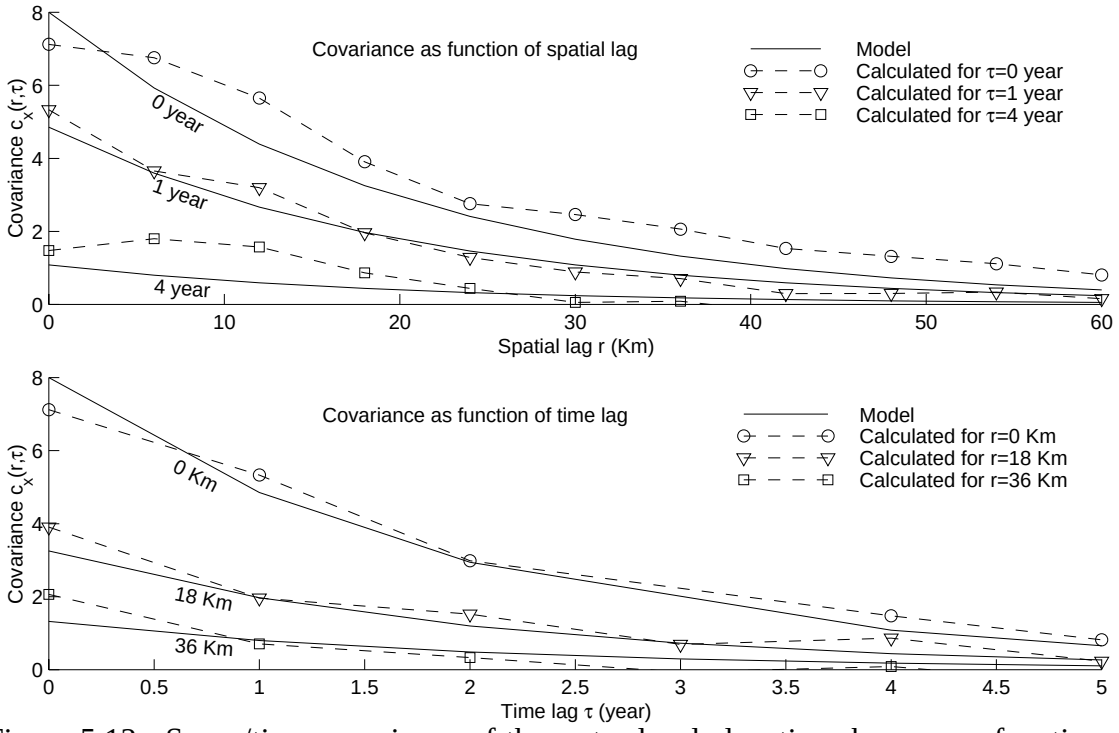

Figure 5.12: Space/time covariance of the water-level elevation shown as a function of spatial lag r (top) and time lag τ (bottom). Markers show calculated covariance values from actual measurements. The plain lines show the fitted covariance model.

as a function of the spatial and temporal lags. The covariance decreases smoothly as r and t increase, and tends to zero for large r and t, thus suggesting that the water-level elevation model of Eq. (5.10) is a reasonable choice. The separable exponential covariance model $c_x(r,\tau) = c_0 \exp[-r / a_r]\exp[-\tau / a_t]$, with sill $c_0 = 8\ ft^2$, spatial range $a_r = 20\ Km$ and temporal range $a_t = 2\ years$ is also plotted in Fig. 5.12. This model offers a reasonably good fit to the estimated covariance values.

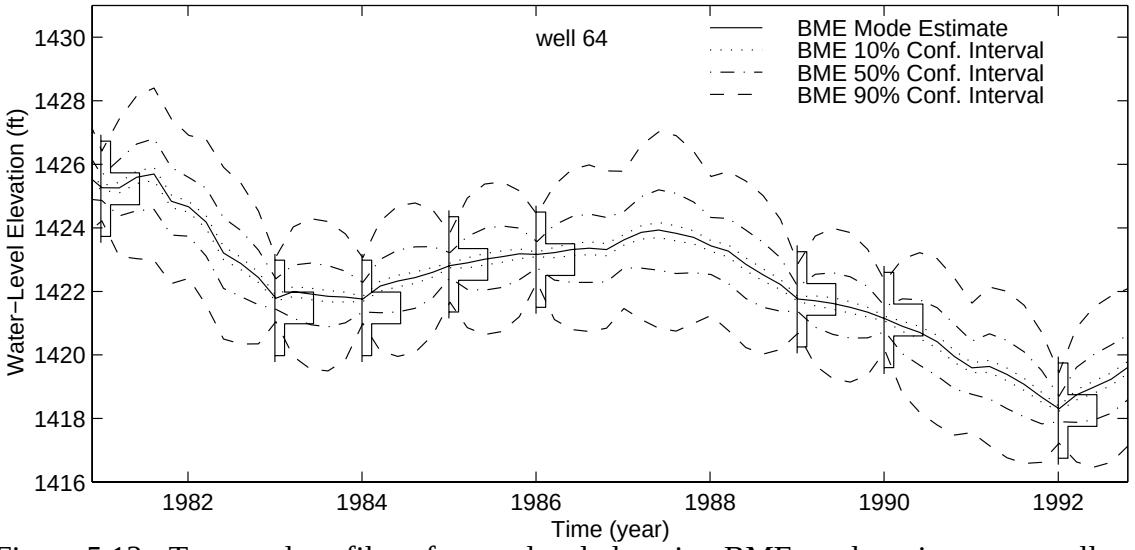

Figure 5.13: Temporal profiles of water-level elevation BME mode estimates at well no. 64. The corresponding BME 10%, 50% and 90% confidence intervals are also shown. The soft probabilistic data are depicted using pulse-shaped curves

The BME approach uses the space/time covariances together with the physical data available (hard and soft), in order to produce spatiotemporal $X(\boldsymbol{p})$-estimates. Then, the corresponding water-level elevations $W(\boldsymbol{p})$ are obtained by adding the (known) spatial mean $\mu(\boldsymbol{s})$ to the $X(\boldsymbol{p})$-estimates. For illustration, the BME mode estimates for well no. 64 are shown in Fig. 5.13 as a function of time (hard and soft data were used in the BME estimation). Also shown are the BME 10%, 50% and 90% confidence intervals (note that estimation at a specific well used space/time data from the same as well as neighboring

wells). For this representative hydrograph the confidence intervals are consistent with the soft probabilistic data (represented by pulse-shaped curves at the observation times); the confidence intervals are wider at times between observations. In Fig. 5.14 we show the BME mode estimates and the 90% confidence intervals of the water-level elevation for a representative set of wells. While the BME mode estimates express the most likely changes

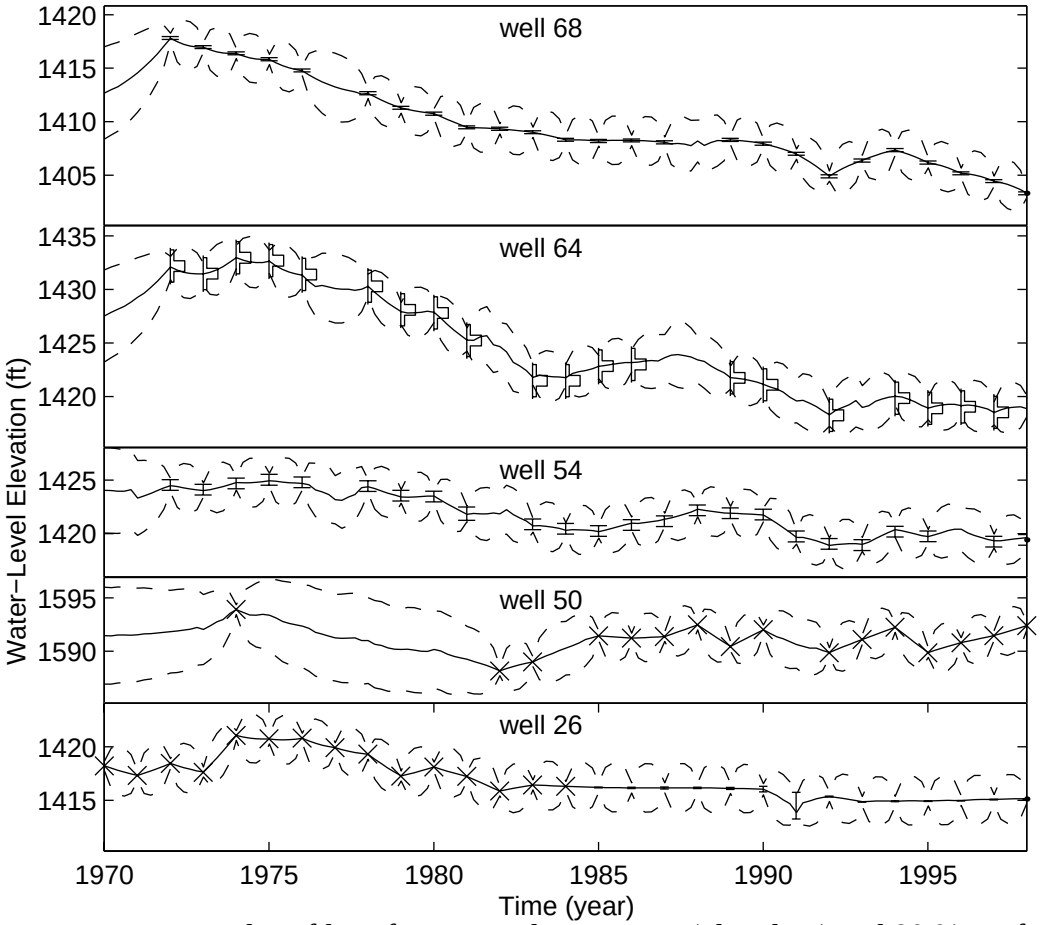

Figure 5.14: Temporal profiles of BME mode estimates (plain line) and 90 % confidence interval (dashed line) of the water-level elevation at selected wells shown. Hard data are shown using $\times$'s; soft data of interval and probabilistic types are depicted using error bars and pulse-shaped curves, respectively.

of the water-level elevation over the time period 1970-1998, the BME 90% intervals provide a good assessment of the associated estimation errors. One can see from these hydrographs that most wells show a decline in the water-level elevation from the late

1970's to the 1988-1992 drought. The water-level elevation reaches its lowest level during 1991-1993, followed by a moderate recovering in the late 1990's. The BME method accounts rigorously for both hard and soft data and provides smaller confidence intervals than kriging methods. These kriging methods lack a rigorous mechanism that would allow them to incorporate many important forms of general knowledge and soft data. Instead, as was demonstrated earlier in this work, most traditional kriging methods derive confidence intervals based on the assumption of a Gaussian posterior pdf, which leads, in general, to more uncertain estimates than BME. For comparison purposes, the 90% confidence intervals obtained using BME and SK are plotted in Fig. 5.15. Note that the SK 90% confidence intervals are at times much wider and, thus, much less informative than the BME 90% confidence intervals.

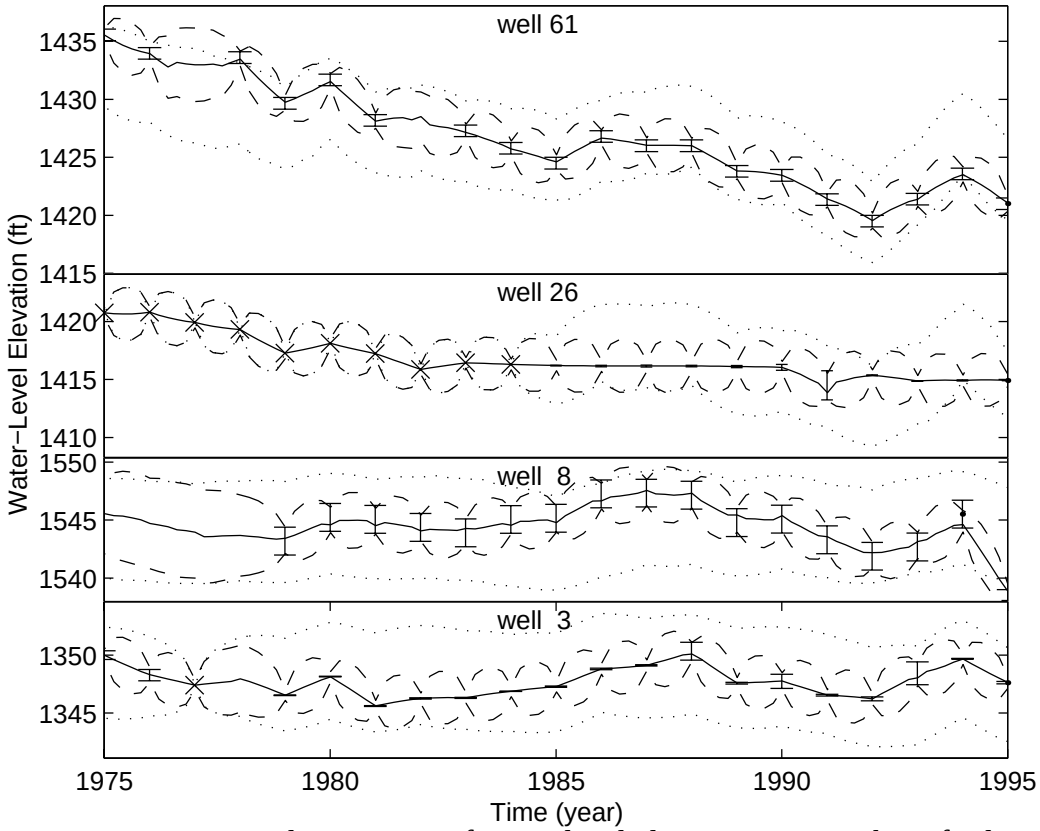

Figure 5.15: BME mode estimates of water-level elevation at a number of selected wells (plain line). Also, 90% confidence intervals obtained from BME (dashed line) and SK (doted line). Hard data are shown using $\times$'s; soft interval data are depicted as error bars.

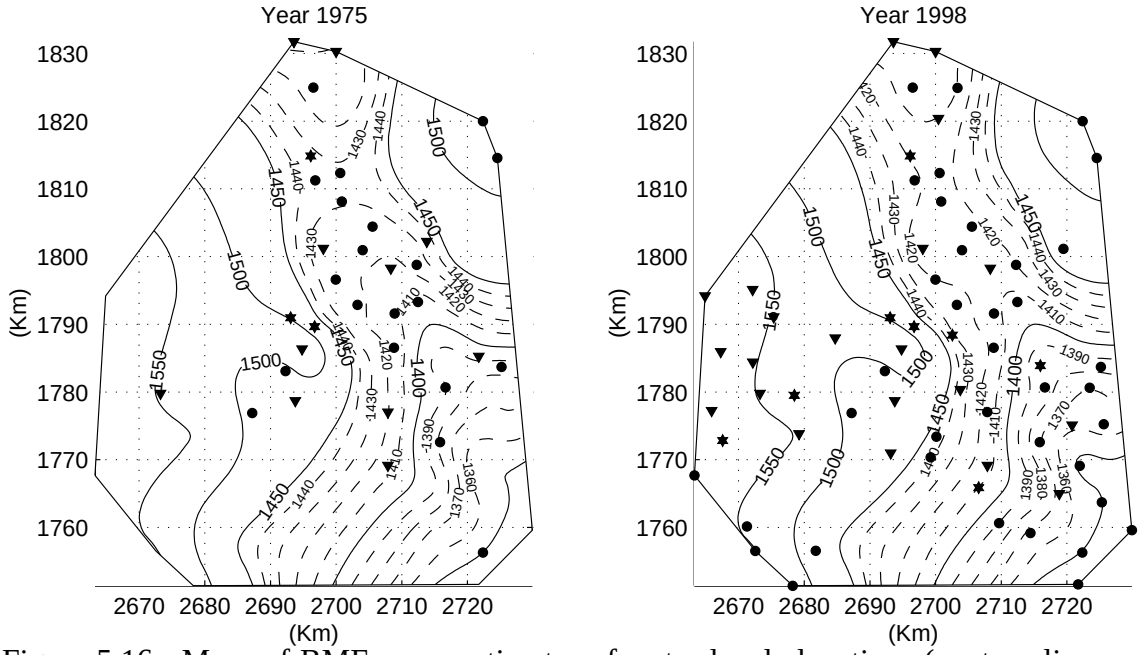

Figure 5.16: Maps of BME mean estimates of water-level elevations (contour lines are annotated in ft) for the years 1975 and 1998. Hard, soft interval and soft probabilistic data are shown with triangles, circles and hexagons, respectively.

BME also produces space/time mean estimates of water-level elevation, which possess high estimation accuracy and are physically informative. Fig. 5.16 shows the BME mean estimates of the water-level elevations during the years 1975 and 1998. Contour lines of water-level elevation are shown (in ft above see-level). The locations where hard data points, soft interval data points, and soft probabilistic data points were available during each one of the years 1975 and 1998 are shown with triangles, circles and hexagons, respectively. In order to construct the map, water-level elevations were estimated on a 40×40 regular grid covering the study area. At each space/time estimation point a local neighborhood of hard and soft data points was used. Each neighborhood includes data points located at spatial distances $< 4a_r$ and temporal lags $< 4a_t$ from the estimation point. (if the local neighborhood includes too many points, then only the 10 hard data points and 5 soft data points most highly correlated to the estimation point are selected). The maps of BME mean estimation of water-level elevation in Fig. 5.16 show

that, while the water-level stayed unchanged at higher grounds, it dropped substantially in the valley along the Little Arkansas river.

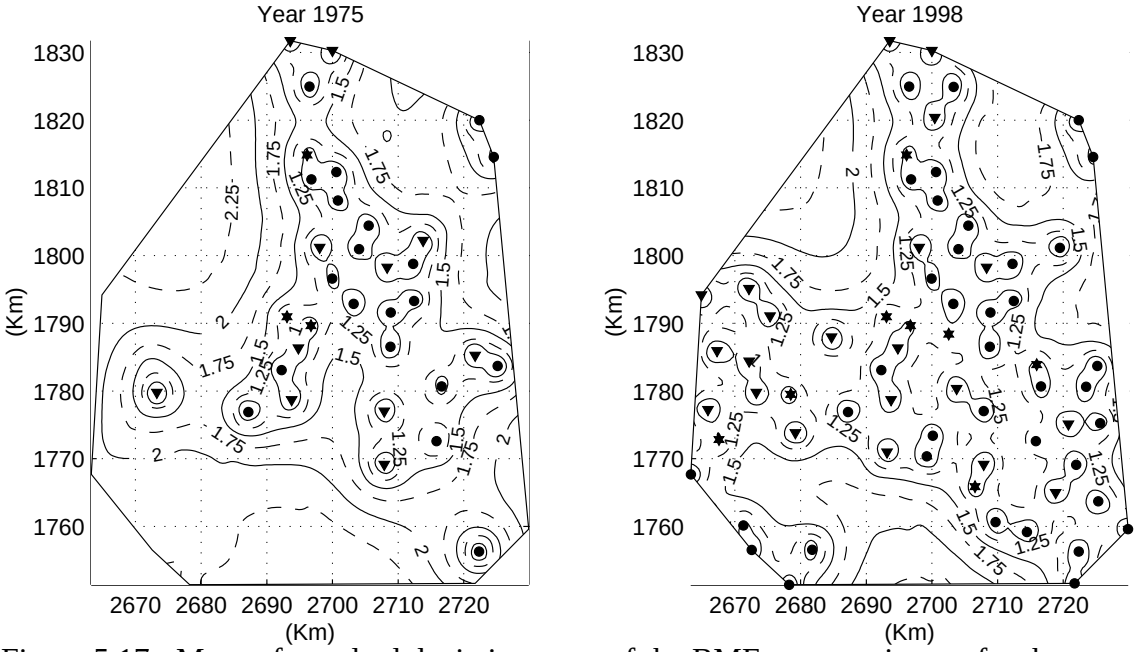

Figure 5.17: Maps of standard deviation error of the BME mean estimates for the water-level elevations (in ft) for the years 1975 and 1998. Hard, soft interval and soft probabilistic data are shown with triangles, circles and hexagons, respectively.

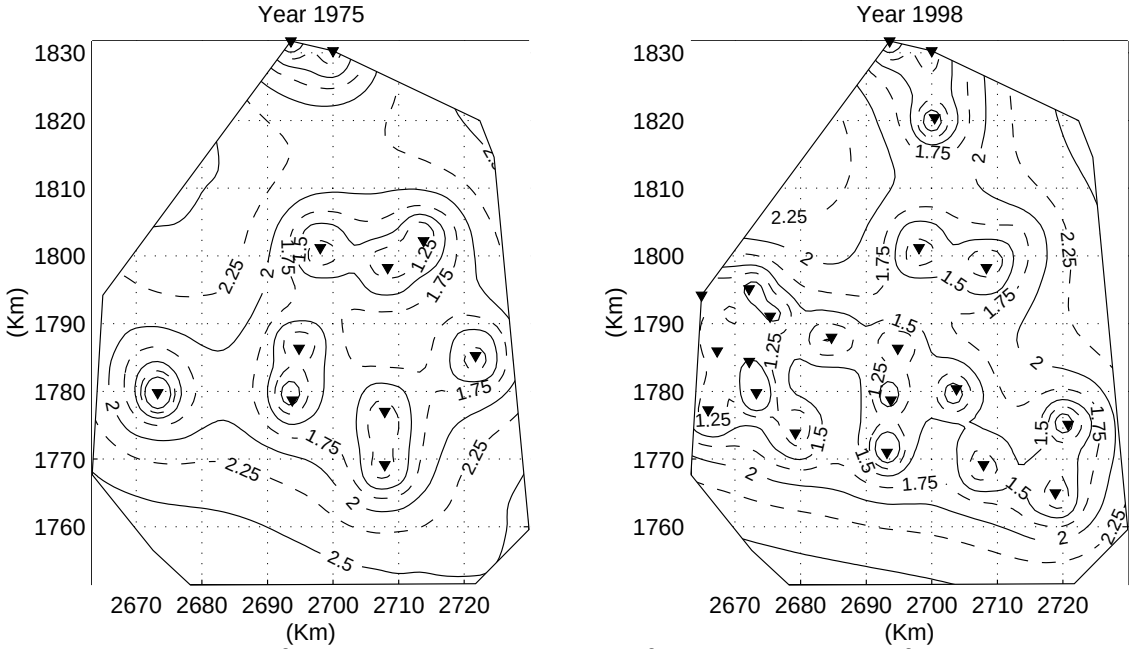

Figure 5.18: Maps of standard deviation error of the SK estimates for the water-level elevations (in ft) for the years 1975 and 1998. Hard data are shown with triangles.

The estimation error for the BME map in Fig. 5.16 is assessed by means of the corresponding estimation error standard deviation (Fig. 5.17). The estimation error standard deviation for the BME map is generally less than 1.5 *ft* in the valley where a significant number of observation wells are located, and increases to about 2 *ft* at higher grounds. As was expected, the estimation error standard deviation calculated using SK reveals larger errors (Fig. 5.18).

The changes in water-level elevations between 1975 and a few selected years are shown in Fig. 5 19. These maps show a continuing depletion during the periods 1975-1980, 1975-1985 and 1975-1992; a modest recovery followed during the period 1992-1998. A zone of water-level decline was developed during the period 1975-1980 in an region generally encompassing the location of the city pumping wells in the north of the study area (Aucott and Myers, 1998). The zone of water-level depletion extended further in the periods ending in 1985 and 1992 to encompass all the well-field, while the maximum water-level decline from 1975 was more than 12 *ft* in 1985 and more than 18 *ft* in 1992. This water-level decline was due to increased water pumping for municipal use, greatly increased agricultural withdrawals, and the combined effect of the 1988-1992 drought (Aucott and Myers, 1998). The effect of ground-water level depletion includes loss of water saturated thickness, increased pumping costs to lift water from greater depths, greater exposure to salt water intrusion from natural and anthropogenic sources to the West. This generally represents a decrease in the water resource available for use. The period 1992-1998 is characterized by some recovery in water-levels due primarily to a decrease in pumping for municipal use. However, as is shown in the maps of water-level change, the extend of the water-level decline since 1975 is still large, with a maximum decline of more than 12 *ft*.

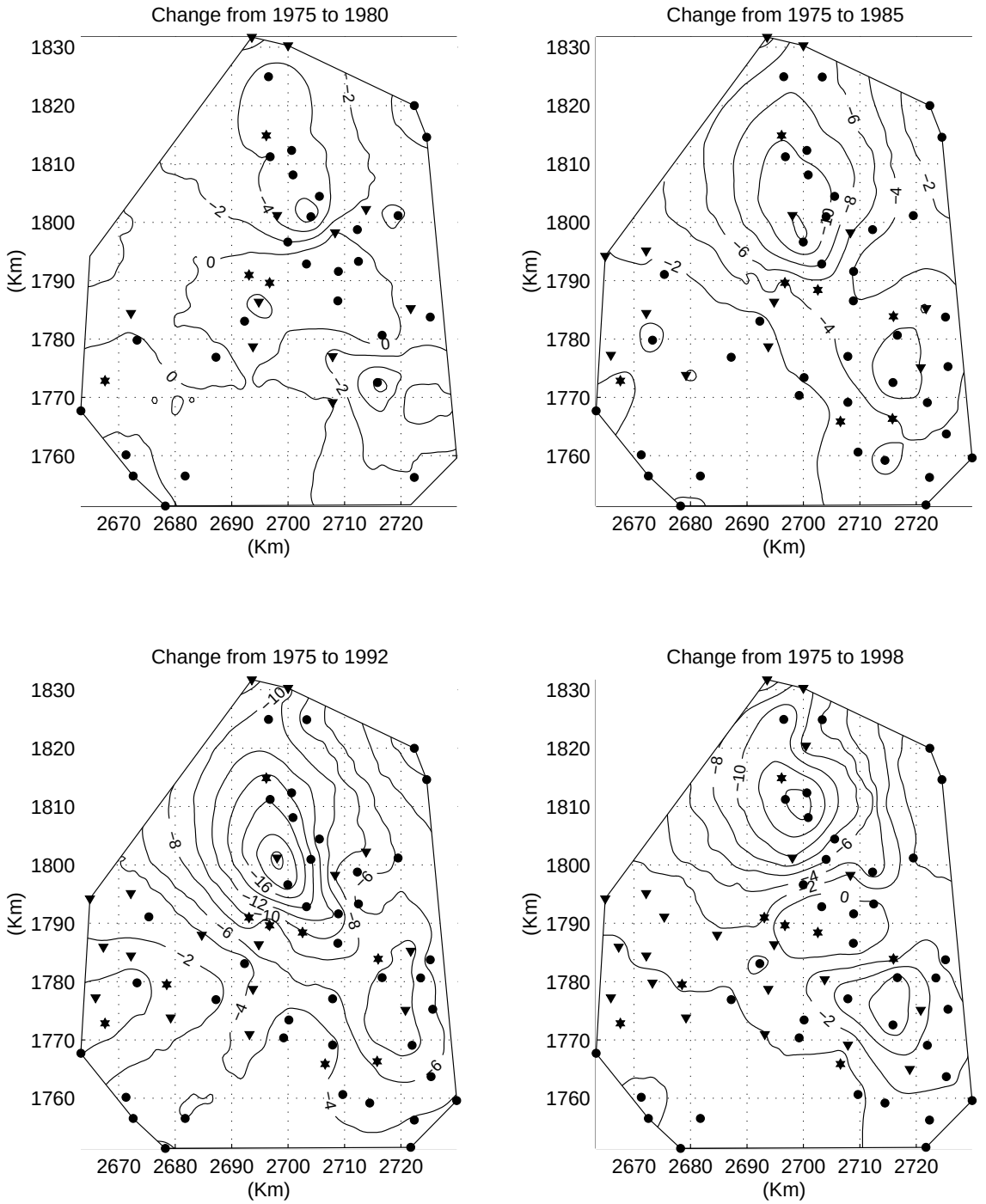

Figure 5.19: Maps of the change in water-level elevation (in *ft*) from 1975 to selected years, obtained using BME mean estimates. Hard, soft interval and soft probabilistic data are shown with triangles, circles and hexagons, respectively.

This case study shows that having accurate and informative estimation tools is valuable in decision making and in the implementation of strategies to improve the water quality of aquifers, such as in the case of the Equus Beds Recharge Demonstration Project.